

AD-A190 327

EMPIRICAL AND HIERARCHICAL BAYES COMPETITORS OF
PRELIMINARY TEST ESTIMATO. (U) PITTSBURGH UNIV PA
CENTER FOR MULTIVARIATE ANALYSIS H GHOSH ET AL. SEP 87
TR-87-34 AFOSR-TR-87-1960 F49620-85-C-0000 F/O 12/3

1/1

UNCLASSIFIED

NL





1.0



1.1



1.25



1.4

1.6



1.8



2.0



2.2



2.5



2.8



3.2



3.6



4.0

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

2

AD-A190 327

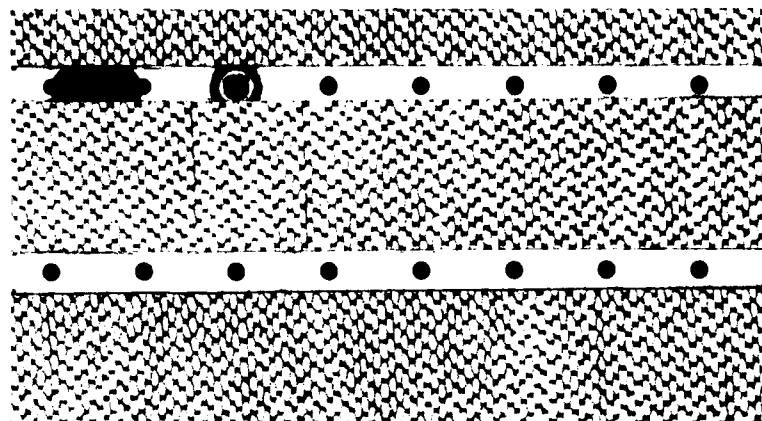
REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 87-1960	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER DTIC FILE COPY
4. TITLE (and Subtitle) Empirical and Hierarchical Bayes competitors of preliminary test estimators in two sample problems		5. TYPE OF REPORT & PERIOD COVERED <i>Journal</i> Technical - September 1987
AUTHOR(s) Malay Ghosh and Bimal K. Sinha		6. PERFORMING ORG. REPORT NUMBER 87-34
PERFORMING ORGANIZATION NAME AND ADDRESS Center for Multivariate Analysis University of Pittsburgh Pittsburgh, PA 15260		8. CONTRACT OR GRANT NUMBER(s) F49620-85-C-0008
CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research Dept. of the Air Force Bolling Air Force Base, DC 20332		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS <i>WICOF 3304 R5</i>
MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) <i>Source as 11</i>		12. REPORT DATE September 1987
		13. NUMBER OF PAGES 27
		15. SECURITY CLASS. (of this report) unclassified
		16. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) DTIC ELECTE S D JAN 15 1988 H		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)		

DD FORM 1 JAN 73 1473

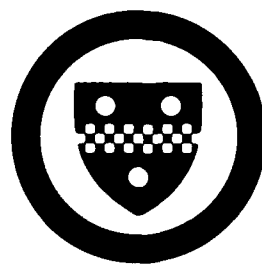
Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

APOSR-TR. 87-1960



Center for Multivariate Analysis
University of Pittsburgh



EMPIRICAL AND HIERARCHICAL BAYES
COMPETITORS OF PRELIMINARY TEST
ESTIMATORS IN TWO SAMPLE PROBLEMS

Malay Ghosh
University of Florida

and

Bimal K. Sinha *
Center for Multivariate Analysis
University of Pittsburgh
and
University of Maryland Baltimore County

September 1987

Technical Report No. 87-34



Accession For

NTIS GRA&I ☒

DTIC TAB

Unannounced

Justification

Center for Multivariate Analysis
Fifth Floor, Thackeray Hall
University of Pittsburgh
Pittsburgh, PA 15260

For

Information

Reference

Number

Accession

Number

Accession

Number

Accession

Number

Accession

Number

Accession

Number

Accession

Number

Accession

Number

Accession

Number

Accession

Number

Accession

Number

*Research partially supported by the Air Force Office of Scientific Research (AFOSR) under contract #F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

**EMPIRICAL AND HIERARCHICAL BAYES
COMPETITORS OF
PRELIMINARY TEST ESTIMATORS IN TWO
SAMPLE PROBLEMS**

Malay Ghosh*
University of Florida

Bimal K. Sinha†
University of Maryland Baltimore County
and
Center for Multivariate Analysis,
University of Pittsburgh

*Research partially supported by NSF Grant Number DMS 8600666

†Research partially supported by Air Force Office of Scientific Research (AFOSR)
under contract F49620-85-C-0008.

1. Introduction

Suppose in a laboratory, say Laboratory I, a certain instrument is designed to measure several characteristics, and a number of vector-valued measurements is recorded. Our objective is to estimate the unknown population mean. It is known, however, that a similar instrument is used in another laboratory, say Laboratory II for the same purpose, and a number of observations is recorded from the second instrument. It is also suspected that the two population means are equal, in which case, observations recorded in Laboratory II can possibly be used effectively together with those in Laboratory I for estimating the population mean of the first instrument. Thus, the question that naturally arises is whether one should use the sample mean from Laboratory I or the pooled mean from the two laboratories.

In problems of this type what is normally sought is a compromise estimator which leans more towards the pooled sample mean when the null hypothesis of the equality of the two population means is accepted, and towards the sample mean from Laboratory I when such a hypothesis is rejected.

A very popular way to achieve this compromise is to use a preliminary test estimator (PTE) which uses the pooled mean when the null hypothesis is accepted at a desired level of significance, and uses the sample mean from Laboratory I when opposite is the case. For an excellent review of PTE's, see Bancroft and Han (1981). It is known, though, in other situations that a PTE is typically not a minimax estimator, and estimators with uniformly smaller mean squared error (MSE) than the PTE can often be produced (see for example Sclove et al (1972)). Moreover, the degree of evidence for or against the null hypothesis is not reflected in the PTE.

In this paper, we propose instead an empirical Bayes (EB) estimator which achieves the intended compromise. Such an EB estimator is quite often a weighted average of the pooled mean and the first sample mean. The weights are adaptively determined from the data in such a way that larger the value of the usual F statistic used for testing the equality of the two population means, the smaller is the weight attached towards the pooled sample mean. Thus, unlike the PTE, the EB estimator incorporates the degree of evidence for or against the null hypothesis in a very natural way. Also, unlike a subjective Bayes estimator, the EB estimator is quite robust (with respect to its frequentist or Bayesian risk) against a wide class of priors.

Section 2 motivates the EB estimator, and its Bayesian properties are discussed in this Section. Among other things, it is shown that the EB estimator has uniformly smaller Bayes risk than the first sample mean. In Section 3, the estimators are compared in terms of their frequentist risks, and sufficient conditions under which an EB estimator dominates the first sample mean are given. Also, in this section, a modified EB estimator is proposed, and sufficient conditions under which it dominates the PTE are given. Finally, in Section 5, a hierarchical Bayes approach is proposed as an alternative to EB estimators. It has recently

come to our attention that Saleh and Ahmed (1987) have considered estimation of μ_1 under the loss $L(\delta, \mu_1) = (\delta - \mu_1)' V^{-1} (\delta - \mu_1)$, assuming $V_1 = V_2 = V$ unknown, and proposed the shrinkage estimator $\bar{x}_1 + \frac{n_2 c}{n_1 + n_2} (\bar{x}_2 - \bar{x}_1) \cdot \frac{n}{J_n^2}$ where $J_n^2 = \frac{n_1 n_2}{n_1 + n_2} (\bar{x}_2 - \bar{x}_1)' S^{-1} (\bar{x}_2 - \bar{x}_1)$, nS = pooled sum of squares and products matrix, $n = n_1 + n_2 - 2$, and $0 < c < \frac{2(p-2)}{n_1 + n_2 - p + 1}$. A comparison of the risk of the above estimator with those of the PTE as well as \bar{x}_1 and $\frac{n_1 \bar{x}_1 + n_2 \bar{x}_2}{n_1 + n_2}$ is also undertaken by the above authors.

2. The EB Estimator and its Bayesian Properties

Let $\mathbf{X}_{1i} (i = 1, \dots, n_1)$ and $\mathbf{X}_{2i} (i = 1, \dots, n_2)$ be independent $p (\geq 3)$ -dimensional random vectors, where \mathbf{X}_{1i} 's are iid $N_p(\mu_1, \sigma^2 \mathbf{V}_1)$, while \mathbf{X}_{2i} 's are iid $N_p(\mu_2, \sigma^2 \mathbf{V}_2)$. In the above $\mu_1 \in R^p$, $\mu_2 \in R^p$ and $\sigma^2 (> 0)$ are unknown, but \mathbf{V}_1 and \mathbf{V}_2 are known $p \times p$ p.d. matrices. Our goal is to estimate μ_1 .

In order to motivate the EB estimator, we need find first a Bayes procedure. It is immediate that the minimal sufficient statistic for (μ_1, μ_2, σ^2) is $(\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2, \text{tr}(\mathbf{V}_1^{-1} \mathbf{S}_1 + \mathbf{V}_2^{-1} \mathbf{S}_2))$, where $\bar{\mathbf{X}}_j = n_j^{-1} \sum_{i=1}^{n_j} \mathbf{X}_{ji} (j = 1, 2)$ and $\mathbf{S}_j = \sum_{i=1}^{n_j} (\mathbf{X}_{ji} - \bar{\mathbf{X}}_j)(\mathbf{X}_{ji} - \bar{\mathbf{X}}_j)^T, j = 1, 2$. Note also that $\bar{\mathbf{X}}_j \sim N_p(\mu_j, \sigma^2 n_j^{-1} \mathbf{V}_j) (j = 1, 2)$, while $\text{tr}(\mathbf{V}_1^{-1} \mathbf{S}_1 + \mathbf{V}_2^{-1} \mathbf{S}_2) \sim \sigma^2 \chi_{(n_1 + n_2 - 2)p}^2$.

In a Bayesian frame work, the above is treated as a conditinal distribution given μ_1 and μ_2 . We use the independent $N_p(\nu, \tau^2 n_1^{-1} \mathbf{V}_1)$ and $N_p(\nu, \tau^2 n_2^{-1} \mathbf{V}_2)$ priors for μ_1 and μ_2 , that is the prior variance - covariance matrix is proportional to the variance-covariance matrix of the corresponding sample mean. The suspicion that μ_1 and μ_2 may be equal is reflected in the choice of a priori common mean ν . For a related prior in the general regression model, see Ghosh et al (1987).

In order to find the posterior distribution of $\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$, first note that conditional on μ_1 and μ_2 , $\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2, \mathbf{S}_1$ and \mathbf{S}_2 are mutually independent, and the distributions of $\mathbf{S}_1, \mathbf{S}_2$ do not depend on μ_1 and μ_2 . Hence, we can restrict ourselves to the conditional distributions of $\bar{\mathbf{X}}_j$'s given μ_j 's. Also, since μ_1 and μ_2 have independent normal priors, standard calculations yield that μ_1 and μ_2 given $\bar{\mathbf{X}}_1$ and $\bar{\mathbf{X}}_2$ have independent posterior distributions with

$$\mu_j | \bar{\mathbf{X}}_j = \bar{\mathbf{x}}_j \sim N_p((1 - B)\bar{\mathbf{x}}_j + B\nu, \sigma^2(1 - B)n_j^{-1}\mathbf{V}_j) \quad (2.1)$$

$j = 1, 2$ where $B = \sigma^2 / (\sigma^2 + \tau^2)$. Now, using the loss

$$L(\mu_1, \mathbf{a}) = \sigma^{-2} (\mathbf{a} - \mu_1)^T \mathbf{Q} (\mathbf{a} - \mu_1) \quad (2.2)$$

for estimating μ_1 by \mathbf{a} (\mathbf{Q} being a known p.d. weight matrix), the Bayes estimator of μ_1 is

$$e_B(\bar{\mathbf{X}}_1) = (1 - B)\bar{\mathbf{X}}_1 + B\nu. \quad (2.3)$$

Note that the Bayes estimator does not depend on the choice of \mathbf{Q} . The multiplier σ^{-2} is used in the loss because that makes $\bar{\mathbf{X}}_1$ a minimax estimator of μ_1 with constant risk not depending on any unknown parameter.

In order to find an EB estimator of μ_1 , we estimate the unknown parameters B and ν in (2.3) from the marginal distributions of $\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2$ and $\text{tr}(\mathbf{V}_1^{-1} \mathbf{S}_1 + \mathbf{V}_2^{-1} \mathbf{S}_2)$. Note that marginally $\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2$ and $\text{tr}(\mathbf{V}_1^{-1} \mathbf{S}_1 + \mathbf{V}_2^{-1} \mathbf{S}_2)$ are mutually independent with $\bar{\mathbf{X}}_j \sim N_p(\nu, n_j^{-1}(\sigma^2 + \tau^2)\mathbf{V}_j), (j = 1, 2)$ and $\text{tr}(\mathbf{V}_1^{-1} \mathbf{S}_1 + \mathbf{V}_2^{-1} \mathbf{S}_2) \sim \sigma^2 \chi_{(n_1 + n_2 - 2)p}^2$. Hence the

complete sufficient statistic for (ν, τ^2, σ^2) based on this marginal distribution is $(\mathbf{W}, Z, \text{tr}(\mathbf{V}_1^{-1}\mathbf{S}_1 + \mathbf{V}_2^{-1}\mathbf{S}_2))$ where $\mathbf{W} = (n_1\mathbf{V}_1^{-1} + n_2\mathbf{V}_2^{-1})^{-1}(n_1\mathbf{V}_1^{-1}\bar{\mathbf{X}}_1 + n_2\mathbf{V}_2^{-1}\bar{\mathbf{X}}_2)$ is the pooled sample mean, $Z = \mathbf{Y}^T(n_1^{-1}\mathbf{V}_1 + n_2^{-1}\mathbf{V}_2)^{-1}\mathbf{Y}$ and $\mathbf{Y} = \bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2$. Also, marginally, $\mathbf{W} \sim N_p(\nu, (\sigma^2 + \tau^2)(n_1\mathbf{V}_1^{-1} + n_2\mathbf{V}_2^{-1}))$, $\mathbf{Y} \sim N_p(0, (n_1^{-1}\mathbf{V}_1 + n_2^{-1}\mathbf{V}_2)(\sigma^2 + \tau^2))$ and $\text{tr}(\mathbf{V}_1^{-1}\mathbf{S}_1 + \mathbf{V}_2^{-1}\mathbf{S}_2) \sim \sigma^2\chi_{(n_1+n_2-2)p}^2$. Hence, the UMVUE of ν is \mathbf{W} , while the UMVUE of $(\sigma^2 + \tau^2)^{-1}$ is $(p-2)/(\mathbf{Y}^T(n_1^{-1}\mathbf{V}_1 + n_2^{-1}\mathbf{V}_2)^{-1}\mathbf{Y})$. The last assertion follows since $\mathbf{Y}^T(n_1^{-1}\mathbf{V}_1 + n_2^{-1}\mathbf{V}_2)^{-1}\mathbf{Y} \sim (\sigma^2 + \tau^2)\chi_p^2$. Moreover since $\text{tr}(\mathbf{V}_1^{-1}\mathbf{S}_1 + \mathbf{V}_2^{-1}\mathbf{S}_2) \sim \sigma^2\chi_{(n_1+n_2-2)p}^2$, the best scale invariant estimator of σ^2 is $((n_1 + n_2 - 2)p + 2)^{-1}\text{tr}(\mathbf{V}_1^{-1}\mathbf{S}_1 + \mathbf{V}_2^{-1}\mathbf{S}_2)$. Substituting these estimators for $\nu, (\sigma^2 + \tau^2)^{-1}$ and σ^2 in (2.3), one gets the EB estimator of μ_1 as

$$\mathbf{e}_{EB}(\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2, \mathbf{S}_1, \mathbf{S}_2) = (1 - \hat{B})\bar{\mathbf{X}}_1 + \hat{B}\mathbf{W} = \mathbf{W} + (1 - \hat{B})(\bar{\mathbf{X}}_1 - \mathbf{W}) \quad (2.4)$$

where

$$\hat{B} = \frac{(p-2)\text{tr}(\mathbf{V}_1^{-1}\mathbf{S}_1 + \mathbf{V}_2^{-1}\mathbf{S}_2)}{((n_1 + n_2 - 2)p + 2)\mathbf{Y}^T(n_1^{-1}\mathbf{V}_1 + n_2^{-1}\mathbf{V}_2)^{-1}\mathbf{Y}} \quad (2.5)$$

Remark 2.1 Note that $0 < B < 1$, while the estimator \hat{B} though positive can values exceeding one. Accordingly, for practical purposes, one proposes the positive part EB estimator

$$\mathbf{e}_{EB}^+(\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2, \mathbf{S}_1, \mathbf{S}_2) = \mathbf{W} + (1 - \hat{B})^+(\bar{\mathbf{X}}_1 - \mathbf{W}). \quad (2.6)$$

of μ_1 , where $a^+ = \max(a, 0)$. For simplicity of exposition, in the remainder of this section, we shall, however, work with \mathbf{e}_{EB} rather than \mathbf{e}_{EB}^+ .

A question that naturally arises is why this particular method of estimation is used for estimating the prior parameters. We shall answer the question by proving the "optimality" of \mathbf{e}_{EB} within the class of estimators

$$\delta_c(\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2, \mathbf{S}_1, \mathbf{S}_2) = \mathbf{W} + (1 - \frac{c\text{tr}(\mathbf{V}_1^{-1}\mathbf{S}_1 + \mathbf{V}_2^{-1}\mathbf{S}_2)}{((n_1 + n_2 - 2)p - 2)\mathbf{Y}^T(n_1^{-1}\mathbf{V}_1 + n_2^{-1}\mathbf{V}_2)^{-1}\mathbf{Y}})(\bar{\mathbf{X}}_1 - \mathbf{W}), \quad (2.7)$$

where $c(> 0)$ is a constant. Note that $\mathbf{e}_{EB} = \delta_{p-2}$.

Theorem 2.1 The Bayes risk of δ_c under the assumed prior (say ξ) and the loss (2.2) is given by

$$\begin{aligned} r(\xi, \delta_c) &= (1 - B)n_1^{-1}\text{tr}(Q\mathbf{V}_1) + B\text{tr}(Q(n_1\mathbf{V}_1^{-1} + n_2\mathbf{V}_2^{-1})^{-1}) \\ &\quad + B\text{tr}(Q\Lambda(n_1^{-1}\mathbf{V}_1 + n_2^{-1}\mathbf{V}_2)\Lambda^T) \\ &\quad \cdot \left[\frac{c^2(n_1 + n_2 - 2)}{\{(n_1 + n_2 - 2)p + 2\}(p - 2)} - \frac{2c(n_1 + n_2 - 2)}{(n_1 + n_2 - 2)p + 2} + 1 \right] \end{aligned} \quad (2.8)$$

where Λ is defined following (2.11) below.

Proof:

The second part of the theorem follows immediately from (2.8). To prove the first part, write

$$r(\xi, \delta_c) = r(\xi, \mathbf{e}_B) + \sigma^{-2} E[(\mathbf{e}_B - \delta_c)^T Q(\mathbf{e}_B - \delta_c)] \quad (2.9)$$

Note from (2.1) to (2.3) that

$$r(\xi, \mathbf{e}_B) = (1 - B)n_1^{-1} \text{tr}(QV_1) \quad (2.10)$$

Also, writing $\hat{B}_c = \text{ctr}(V_1^{-1}S_1 + V_2^{-1}S_2) / \{(n_1 + n_2 - 2)p + 2\} Y^T (n_1^{-1}V_1 + n_2^{-1}V_2)^{-1} Y$, one gets

$$\begin{aligned} \mathbf{e}_B - \delta_c &= (1 - B)\bar{X}_1 + B\nu - \mathbf{W} - (1 - \hat{B}_c)(\bar{X}_1 - \mathbf{W}) \\ &= -B(\mathbf{W} - \nu) + (\hat{B}_c - B)(\bar{X}_1 - \mathbf{W}) \\ &= -B(\mathbf{W} - \nu) + (\hat{B}_c - B)\Lambda Y, \\ \Lambda &= (n_1 V_1^{-1} + n_2 V_2^{-1})^{-1} n_2 V_2^{-1}. \end{aligned} \quad (2.11)$$

Next using the independence of \mathbf{W} and $(Y, \text{tr}(V_1^{-1}S_1 + V_2^{-1}S_2))$ and the facts that $E(\mathbf{W}) = \nu$, $\text{Var}(\mathbf{W}) = (\sigma^2 + \tau^2)(n_1 V_1^{-1} + n_2 V_2^{-1})^{-1} = \sigma^2 B^{-1}(n_1 V_1^{-1} + n_2 V_2^{-1})^{-1}$, one gets

$$\begin{aligned} E[(\mathbf{e}_B - \delta_c)^T Q(\mathbf{e}_B - \delta_c)] &= B^2 E[(\mathbf{W} - \nu)^T Q(\mathbf{W} - \nu)] + E[(\hat{B}_c - B)^2 Y^T \Lambda^T Q \Lambda Y] \\ &= \sigma^2 B \text{tr}\{Q(n_1 V_1^{-1} + n_2 V_2^{-1})^{-1}\} + E[(\hat{B}_c - B)^2 Y^T \Lambda^T Q \Lambda Y] \end{aligned} \quad (2.12)$$

Now we find

$$\begin{aligned} E[(\hat{B}_c - B)^2 Y^T \Lambda^T Q \Lambda Y] &= E\left[\frac{\sigma^2 \{\text{tr}(V_1^{-1}S_1 + V_2^{-1}S_2)\}^2}{\{(n_1 + n_2 - 2)p + 2\}^2 \{Y^T (n_1^{-1}V_1 + n_2^{-1}V_2)^{-1} Y\}^2} (Y^T \Lambda^T Q \Lambda Y) \right. \\ &\quad \left. - \frac{2Bc \text{tr}(V_1^{-1}S_1 + V_2^{-1}S_2)}{\{(n_1 + n_2 - 2)p + 2\} \{Y^T (n_1^{-1}V_1 + n_2^{-1}V_2)^{-1} Y\}} (Y^T \Lambda^T Q \Lambda Y) \right. \\ &\quad \left. + B^2 (Y^T \Lambda^T Q \Lambda Y) \right] \end{aligned} \quad (2.13)$$

Using the independence of Y and $\text{tr}(V_1^{-1}S_1 + V_2^{-1}S_2)$ along with the fact that $\text{tr}(V_1^{-1}S_1 + V_2^{-1}S_2) \sim \sigma^2 \chi_{(n_1 + n_2 - 2)p}^2$, it follows that the right hand side of (2.13)

$$\begin{aligned}
&= E \left[\frac{c^2 \sigma^4 (n_1 + n_2 - 2)p}{(n_1 + n_2 - 2)p + 2} \cdot \frac{\mathbf{Y}^T \Lambda^T Q \Lambda \mathbf{Y}}{\{\mathbf{Y}^T (n_1^{-1} V_1 + n_2^{-1} V_2)^{-1} \mathbf{Y}\}^2} \right. \\
&\quad \left. - \frac{2Bc\sigma^2 (n_1 + n_2 - 2)p}{(n_1 + n_2 - 2)p + 2} \cdot \frac{\mathbf{Y}^T \Lambda^T Q \Lambda \mathbf{Y}}{\{\mathbf{Y}^T (n_1^{-1} V_1 + n_2^{-1} V_2)^{-1} \mathbf{Y}\}} + B^2 (\mathbf{Y}^T \Lambda^T Q \Lambda \mathbf{Y}) \right]
\end{aligned} \quad (2.14)$$

Next observe that $\mathbf{Y}^T (n_1^{-1} V_1 + n_2^{-1} V_2)^{-1} \mathbf{Y}$ is a function of the complete sufficient statistic while $(\mathbf{Y}^T \Lambda^T Q \Lambda \mathbf{Y}) / (\mathbf{Y}^T (n_1^{-1} V_1 + n_2^{-1} V_2)^{-1} \mathbf{Y})$ is ancillary. Now using Basu's theorem (or Lemma 1 of Ghosh et al (1987)) along with $E(\mathbf{Y}^T \Lambda^T Q \Lambda \mathbf{Y}) = (\sigma^2 + \tau^2) \text{tr}(Q \Lambda (n_1^{-1} V_1 + n_2^{-1} V_2) \Lambda^T)$, $E(\mathbf{Y}^T (n_1^{-1} V_1 + n_2^{-1} V_2)^{-1} \mathbf{Y}) = p(\sigma^2 + \tau^2)$, and $E(\mathbf{Y}^T (n_1^{-1} V_1 + n_2^{-1} V_2)^{-1} \mathbf{Y})^{-1} = (\sigma^2 + \tau^2)^{-1} (p - 2)^{-1}$, it follows that the right hand side of (2.14)

$$\begin{aligned}
&= \frac{c^2 \sigma^2 B (n_1 + n_2 - 2) p \text{tr}(Q \Lambda (n_1^{-1} V_1 + n_2^{-1} V_2) \Lambda^T)}{\{(n_1 + n_2 - 2)p + 2\} p (p - 2)} \\
&\quad - \frac{2c \sigma^2 B (n_1 + n_2 - 2) p \text{tr}(Q \Lambda (n_1^{-1} V_1 + n_2^{-1} V_2) \Lambda^T)}{\{(n_1 + n_2 - 2)p + 2\} p} \\
&\quad + \sigma^2 B \text{tr}(Q \Lambda (n_1^{-1} V_1 + n_2^{-1} V_2) \Lambda^T)
\end{aligned} \quad (2.15)$$

It follows from (2.12) - (2.15) that

$$\begin{aligned}
&E[(\mathbf{e}_B - \delta_c)^T Q (\mathbf{e}_B - \delta_c)] \\
&= \sigma^2 B \text{tr}(Q (n_1 V_1^{-1} + n_2 V_2^{-1})^{-1}) + \sigma^2 B \text{tr}(Q \Lambda (n_1^{-1} V_1 + n_2^{-1} V_2) \Lambda^T) \\
&\quad \cdot \left[\frac{c^2 (n_1 + n_2 - 2)}{\{(n_1 + n_2 - 2)p + 2\} (p - 2)} - \frac{2c (n_1 + n_2 - 2)}{(n_1 + n_2 - 2)p + 2} + 1 \right]
\end{aligned} \quad (2.16)$$

The proof of the theorem is complete from (2.9), (2.10) and (2.16).

Next we compare the Bayes risks of \mathbf{e}_{EB} and $\bar{\mathbf{X}}_1$. Note that $\bar{\mathbf{X}}_1$ has constant risk, and hence constant Bayes risk (under any prior) $\sigma^2 n_1^{-1} \text{tr}(Q V_1)$. Rather than comparing the Bayes risks of \mathbf{e}_{EB} and $\bar{\mathbf{X}}_1$ directly, we find it convenient to introduce the notion of Relative Savings Loss (RSL) as in Efron and Morris (1973).

For any estimator \mathbf{e} of μ_1 , the RSL of \mathbf{e}_{EB} with respect to \mathbf{e} (under the prior ξ) is defined as

$$\begin{aligned}
RSL(\xi; \mathbf{e}_{EB}, \mathbf{e}) &= [r(\xi, \mathbf{e}_{EB}) - r(\xi, \mathbf{e})] / [r(\xi, \mathbf{e}) - r(\xi, \mathbf{e}_B)] \\
&= 1 - [r(\xi, \mathbf{e}) - r(\xi, \mathbf{e}_{EB})] / [r(\xi, \mathbf{e}) - r(\xi, \mathbf{e}_B)]
\end{aligned} \quad (2.17)$$

This is the proportion of the possible Bayes risk improvement over \mathbf{e} that is sacrificed by the use of \mathbf{e}_{EB} rather than the ideal \mathbf{e}_B under the prior ξ . From (2.8) with $c = p - 2$ and (2.10), it follows that

$$\begin{aligned}
& RSL(\xi; \mathbf{e}_{ER}, \tilde{\mathbf{X}}_1) \\
&= [\text{tr}(Q(n_1 \mathbf{V}_1^{-1} + n_2 \mathbf{V}_2^{-1})^{-1}) \\
&+ \text{tr}(Q\Lambda(n_1^{-1}V_1 + n_2^{-1}V_2)\Lambda^T) \left(\frac{2(n_1 + n_2 - 1)}{(n_1 + n_2 - 2)p + 2} \right)] [n_1^{-1} \text{tr}(QV_1)]^{-1} \quad (2.18)
\end{aligned}$$

Note that the above RSL expression does not depend on any unknown parameter. Also, from (2.18), since $2(n_1 + n_2 - 1) < (n_1 + n_2 - 2)p + 2$, and

$$(n_1 \mathbf{V}_1^{-1} + n_2 \mathbf{V}_2^{-1})^{-1} + \Lambda(n_1^{-1}V_1 + n_2^{-1}V_2)\Lambda^T = n_1^{-1}V_1, \quad (2.19)$$

it follows that $RSL(\xi; \mathbf{e}_{ER}, \tilde{\mathbf{X}}_1) < 1$ which is equivalent to $r(\xi, \mathbf{e}_{ER}) < r(\xi, \tilde{\mathbf{X}}_1)$. Thus \mathbf{e}_{ER} has smaller Bayes risk than $\tilde{\mathbf{X}}_1$. A proof of (2.19), follows easily from simultaneous diagonalization of V_1 and V_2 . Alternatively, writing

$$\begin{aligned}
\Lambda &= (n_1 V_1^{-1} + n_2 V_2^{-1})^{-1} n_2 V_2^{-1} = [V_2^{-1} \{n_1 V_2 + n_2 V_1\} V_1^{-1}]^{-1} n_2 V_2^{-1} \\
&= n_1^{-1} V_1 (n_1^{-1} V_1 + n_2^{-1} V_2)^{-1},
\end{aligned}$$

it follows that

$$\begin{aligned}
(n_1 V_1^{-1} + n_2 V_2^{-1})^{-1} + \Lambda(n_1^{-1}V_1 + n_2^{-1}V_2)\Lambda^T &= n_2^{-1}\Lambda V_2 + n_1^{-1}\Lambda V_1 \\
&= \Lambda(n_1^{-1}V_1 + n_2^{-1}V_2) = n_1^{-1}V_1
\end{aligned}$$

which is (2.19).

Finally, in this section, we compare the Bayes risk of \mathbf{e}_{ER} with that of \mathbf{W} . Note that \mathbf{W} has Bayes risk

$$r(\xi, \mathbf{W}) = r(\xi, \mathbf{e}_n) + \sigma^{-2} E[(\mathbf{e}_n - \mathbf{W})^T Q(\mathbf{e}_n - \mathbf{W})] \quad (2.20)$$

Since $\mathbf{e}_n - \mathbf{W} = (1 - B)\tilde{\mathbf{X}}_1 + B\nu - \mathbf{W} = -B(\mathbf{W} - \nu) + (1 - B)(\tilde{\mathbf{X}}_1 - \mathbf{W}) = -B(\mathbf{W} - \nu) + (1 - B)\Lambda \mathbf{Y}$ where Λ is defined following (2.11), using once again the independence of \mathbf{W} and \mathbf{Y} , it follows that

$$\begin{aligned}
& E[(\mathbf{e}_n - \mathbf{W})^T Q(\mathbf{e}_n - \mathbf{W})] \\
&= \sigma^2 B \text{tr}(Q(n_1 \mathbf{V}_1^{-1} + n_2 \mathbf{V}_2^{-1})^{-1}) + (1 - B)^2 E(\mathbf{Y}^T \Lambda^T Q \Lambda \mathbf{Y}) \\
&= \sigma^2 B \text{tr}(Q(n_1 \mathbf{V}_1^{-1} + n_2 \mathbf{V}_2^{-1})^{-1}) + \sigma^2 (1 - B)^2 B^{-1} \text{tr}(Q\Lambda(n_1^{-1}V_1 + n_2^{-1}V_2)\Lambda^T) \quad (2.21)
\end{aligned}$$

Thus from (2.10), (2.20) and (2.21),

$$\begin{aligned} r(\xi, \mathbf{W}) &= n_1^{-1}(1-B)\text{tr}(QV_1) + B\text{tr}(Q(n_1\mathbf{V}_1^{-1} + n_2\mathbf{V}_2^{-1})^{-1}) \\ &\quad + (1-B)^2B^{-1}\text{tr}(Q\Lambda(n_1^{-1}V_1 + n_2^{-1}V_2)\Lambda^T) \end{aligned} \quad (2.22)$$

Finally, from (2.8) with $c = p-2$, (2.10) and (2.22), it follows that $\text{RSL}(\xi; \mathbf{e}_{EB}, \mathbf{W}) = [r(\xi, \mathbf{e}_{EB}) - r(\xi, \mathbf{e}_B)]/[r(\xi, \mathbf{W}) - r(\xi, \mathbf{e}_B)]$

$$= \frac{\text{tr}(Q(n_1\mathbf{V}_1^{-1} + n_2\mathbf{V}_2^{-1})^{-1}) + \frac{2(n_1+n_2-1)}{(n_1+n_2-2)p+2}\text{tr}(Q\Lambda(n_1^{-1}V_1 + n_2^{-1}V_2)\Lambda^T)}{\text{tr}(Q(n_1\mathbf{V}_1^{-1} + n_2\mathbf{V}_2^{-1})^{-1}) + (1-B)^2B^{-1}\text{tr}(Q\Lambda(n_1^{-1}V_1 + n_2^{-1}V_2)\Lambda^T)} \quad (2.23)$$

which is less than one if and only if

$$\{(1-B)/B\}^2 > 2(n_1 + n_2 - 1)/\{(n_1 + n_2 - 2)p + 2\} \quad (2.24)$$

Remark 2.2 The fact that \mathbf{e}_{EB} does not dominate \mathbf{W} uniformly is not at all surprising. If, for example, τ^2 is very small and μ_1 is nearly degenerate at ν , then \mathbf{W} is much closer to ν than \mathbf{e}_{EB} . Indeed, in this case $B = \sigma^2/(\sigma^2 + \tau^2)$ is very close to 1 so that (2.24) cannot hold. However, when $\sigma^2 \leq \tau^2$, then $B \leq \frac{1}{2} \leftrightarrow (1-B)/B \geq 1$ so that (2.24) holds.

3. Minimax Estimation

It is well known that under the loss given in (2.2), \bar{X}_1 is a minimax estimator of μ_1 with constant risk $n_1^{-1} \text{tr}(\mathbf{QV}_1)$. In this section, first we find a class of estimators including e_{EB} as a member which dominates \bar{X}_1 under certain conditions, and then investigate whether e_{EB} satisfies these conditions.

With this end, first write

$$F = (\mathbf{Y}^T(n_1^{-1}\mathbf{V}_1 + n_2^{-1}\mathbf{V}_2)^{-1}\mathbf{Y}) / \{\text{tr}(\mathbf{V}_1^{-1}\mathbf{S}_1 + \mathbf{V}_2^{-1}\mathbf{S}_2) / ((n_1 + n_2 - 2)p + 2)\} \quad (3.1)$$

and consider the class of estimators

$$\mu_1^\phi = \bar{X}_1 - (\phi(F)/F)(\bar{X}_1 - \mathbf{W}) \quad (3.2)$$

for estimating μ_1 . Note that e_{EB} belongs to this class with $\phi(F) = p - 2$. We now compute the frequentist risk of the estimator μ_1^ϕ (i.e. without any reference to the prior ξ). Throughout this section, E denotes expectation conditional on μ_1 and μ_2 .

Theorem 3.1

$$\begin{aligned} & E[(\mu_1^\phi - \mu_1)^T \mathbf{Q}(\mu_1^\phi - \mu_1)] / \sigma^2 \\ &= n_1^{-1} \text{tr}(\mathbf{QV}_1) - 2E \left[\frac{\phi(F)}{F} \text{tr}(\mathbf{\Lambda}^T \mathbf{Q} \mathbf{\Lambda} \mathbf{V}) + 2(\phi'(F) - \frac{\phi(F)}{F}) \frac{\mathbf{Y}^T \mathbf{\Lambda}^T \mathbf{Q} \mathbf{\Lambda} \mathbf{Y}}{\mathbf{Y}^T \mathbf{V}^{-1} \mathbf{Y}} \right] \\ &+ \sigma^{-2} E \left[\frac{\phi^2(F)}{F^2} \mathbf{Y}^T \mathbf{\Lambda}^T \mathbf{Q} \mathbf{\Lambda} \mathbf{Y} \right] \end{aligned} \quad (3.3)$$

where $\mathbf{V} = n_1^{-1}\mathbf{V}_1 + n_2^{-1}\mathbf{V}_2$.

Proof: First write

$$\begin{aligned} & E[(\mu_1^\phi - \mu_1)^T \mathbf{Q}(\mu_1^\phi - \mu_1)] \\ &= E[(\bar{X}_1 - \mu_1)^T \mathbf{Q}(\bar{X}_1 - \mu_1)] \\ &- 2(\phi(F)/F) \mathbf{Y}^T \mathbf{\Lambda}^T \mathbf{Q}(\bar{X}_1 - \mu_1) \\ &+ (\phi^2(F)/F^2) \mathbf{Y}^T \mathbf{\Lambda}^T \mathbf{Q} \mathbf{\Lambda} \mathbf{Y} \end{aligned} \quad (3.4)$$

where we have used the fact that $\bar{X}_1 - \mathbf{W} = \mathbf{\Lambda} \mathbf{Y}$. Next writing $\bar{X}_1 = \mathbf{W} + \mathbf{\Lambda} \mathbf{Y}$ and correspondingly $\mu_1 = \mu_* + \mathbf{\Lambda} \mu_0$ where $\mu_* = (n_1 \mathbf{V}_1^{-1} + n_2 \mathbf{V}_2^{-1})^{-1} (n_1 \mathbf{V}_1^{-1} \mu_1 + n_2 \mathbf{V}_2^{-1} \mu_2)$ and $\mu_0 = \mu_1 - \mu_2$, one gets

$$\begin{aligned} & E[(\phi(F)/F) \mathbf{Y}^T \mathbf{\Lambda}^T \mathbf{Q}(\bar{X}_1 - \mu_1)] \\ &= E[(\phi(F)/F) \mathbf{Y}^T \mathbf{\Lambda}^T \mathbf{Q}((\mathbf{W} - \mu_*) + \mathbf{\Lambda}(\mathbf{Y} - \mu_0))] \\ &= E[(\phi(F)/F) \mathbf{Y}^T \mathbf{\Lambda}^T \mathbf{Q} \mathbf{\Lambda}(\mathbf{Y} - \mu_0)] \end{aligned} \quad (3.5)$$

where in the final step of (3.5), one uses the independence of $(\mathbf{Y}, \text{tr}(V_1^{-1}S_1 + V_2^{-1}S_2))$ with \mathbf{W} as well as $E(\mathbf{W}) = \mu_0$. Now since \mathbf{V} is p.d., there exists a nonsingular \mathbf{D} such that $\mathbf{D}^{-1}\mathbf{V}(\mathbf{D}^{-1})^T = \mathbf{I}_p$. Write $\mathbf{Z} = \mathbf{D}^{-1}\mathbf{Y}$ and $\eta_0 = \mathbf{D}^{-1}\mu_0$. Then $\mathbf{Z} \sim N_p(\eta_0, \sigma^2\mathbf{I}_p)$. We rewrite

$$\mathbf{Y}^T \mathbf{\Lambda}^T \mathbf{Q} \mathbf{\Lambda} (\mathbf{Y} - \mu_0) = \mathbf{Z}^T \mathbf{U} (\mathbf{Z} - \eta_0), \quad (3.6)$$

where $\mathbf{U} = ((u_{ij})) = \mathbf{D}^T \mathbf{\Lambda}^T \mathbf{Q} \mathbf{\Lambda} \mathbf{D}$. Also, in terms of \mathbf{Z} , $F = \mathbf{Z}^T \mathbf{Z} / \{\text{tr}(V_1^{-1}S_1 + V_2^{-1}S_2) / ((n_1 + n_2 - 2)p + 2)\}$. Now using Stein's identity (cf. Stein (1981)), the independence of \mathbf{Z} and $\text{tr}(V_1^{-1}S_1 + V_2^{-1}S_2)$, and (3.6),

$$\begin{aligned} & E[(\phi(F)/F) \mathbf{Z}^T \mathbf{U} (\mathbf{Z} - \eta_0)] \\ &= \sigma^2 \sum_{i=1}^p E \left[\frac{\partial}{\partial Z_i} \left\{ \frac{\phi(F)}{F} \cdot \sum_{j=1}^p u_{ij} Z_j \right\} \right] \\ &= \sigma^2 \sum_{i=1}^p E \left[\frac{\phi(F)}{F} u_{ii} + \left\{ \frac{\phi'(F)}{F} - \frac{\phi(F)}{F^2} \right\} \frac{2Z_i \sum_{j=1}^p u_{ij} Z_j}{\{\text{tr}(V_1^{-1}S_1 + V_2^{-1}S_2) / ((n_1 + n_2 - 2)p + 2)\}} \right] \\ &= \sigma^2 E \left[\frac{\phi(F)}{F} \cdot \text{tr}(\mathbf{U}) + 2 \left\{ \frac{\phi'(F)}{F} - \frac{\phi(F)}{F^2} \right\} \cdot \frac{\mathbf{Z}^T \mathbf{U} \mathbf{Z}}{\{\text{tr}(V_1^{-1}S_1 + V_2^{-1}S_2) / ((n_1 + n_2 - 2)p + 2)\}} \right] \\ &= \sigma^2 E \left[\frac{\phi(F)}{F} \text{tr}(\mathbf{\Lambda}^T \mathbf{Q} \mathbf{\Lambda} \mathbf{V}) + 2 \left\{ \frac{\phi'(F)}{F} - \frac{\phi(F)}{F^2} \right\} \cdot F \cdot \frac{\mathbf{Y}^T \mathbf{\Lambda}^T \mathbf{Q} \mathbf{\Lambda} \mathbf{Y}}{\mathbf{Y}^T \mathbf{V}^{-1} \mathbf{Y}} \right] \\ &= \sigma^2 E \left[\frac{\phi(F)}{F} \text{tr}(\mathbf{\Lambda}^T \mathbf{Q} \mathbf{\Lambda} \mathbf{V}) + 2 \left\{ \frac{\phi'(F)}{F} - \frac{\phi(F)}{F^2} \right\} \cdot \frac{\mathbf{Y}^T \mathbf{\Lambda}^T \mathbf{Q} \mathbf{\Lambda} \mathbf{Y}}{\mathbf{Y}^T \mathbf{V}^{-1} \mathbf{Y}} \right] \quad (3.7) \end{aligned}$$

The theorem follows now from (3.3), (3.4) and (3.7).

Next in this section we find an upper bound for $E[(\phi^2(F)/F^2) \mathbf{Y}^T \mathbf{\Lambda}^T \mathbf{Q} \mathbf{\Lambda} \mathbf{Y}]$. We first get the inequality

$$\begin{aligned} & E[(\phi^2(F)/F^2) (\mathbf{Y}^T \mathbf{\Lambda}^T \mathbf{Q} \mathbf{\Lambda} \mathbf{Y})] \\ &= E \left[\frac{\phi^2(F)}{F^2} \cdot F \cdot \frac{\mathbf{Y}^T \mathbf{\Lambda}^T \mathbf{Q} \mathbf{\Lambda} \mathbf{Y}}{\mathbf{Y}^T \mathbf{V}^{-1} \mathbf{Y}} \cdot \frac{\text{tr}(V_1^{-1}S_1 + V_2^{-1}S_2)}{\{(n_1 + n_2 - 2)p + 2\}} \right] \\ &\leq \text{ch}_1(\mathbf{\Lambda}^T \mathbf{Q} \mathbf{\Lambda} \mathbf{V}) E[h^2(F) F \cdot \text{tr}(V_1^{-1}S_1 + V_2^{-1}S_2) / ((n_1 + n_2 - 2)p + 2)] \quad (3.8) \end{aligned}$$

where $\text{ch}_1(\mathbf{\Lambda}^T \mathbf{Q} \mathbf{\Lambda} \mathbf{V})$ denotes the largest eigen value of $\mathbf{\Lambda}^T \mathbf{Q} \mathbf{\Lambda} \mathbf{V}$ and $h(F) = \phi(F)/F$. Next applying (2.18) of Efron and Morris (1976), one gets

$$E[h^2(F) F \text{tr}(V_1^{-1}S_1 + V_2^{-1}S_2) / ((n_1 + n_2 - 2)p + 2)]$$

$$\begin{aligned}
&= E \left[\frac{(n_1 + n_2 - 2)p}{(n_1 + n_2 - 2)p + 2} \cdot h^2(F)F + \frac{2}{(n_1 + n_2 - 2)p + 2} \cdot \frac{\text{tr}(V_1^{-1}S_1 + V_2^{-1}S_2)}{(n_1 + n_2 - 2)p + 2} \right. \\
&\quad \cdot \left. (2h(F)h'(F)F + h^2(F)) \left(-\frac{F}{\text{tr}(V_1^{-1}S_2 + V_2^{-1}S_2)/((n_1 + n_2 - 2)p + 2)} \right) \right] \\
&= \sigma^2 E \left[\frac{(n_1 + n_2 - 2)p}{(n_1 + n_2 - 2)p + 2} \cdot \frac{\phi^2(F)}{F} \right. \\
&\quad \left. - 2 \frac{F}{(n_1 + n_2 - 2)p + 2} \cdot \left\{ 2 \left(\frac{\phi'(F)}{F} - \frac{\phi(F)}{F^2} \right) \phi(F) + \frac{\phi^2(F)}{F^2} \right\} \right] \\
&= \sigma^2 E \left[\frac{\phi^2(F)}{F} - \frac{4}{(n_1 + n_2 - 2)p + 2} \phi(F)\phi'(F) \right] \tag{3.9}
\end{aligned}$$

From (3.8) and (3.9), one gets

$$\begin{aligned}
&E \left[\frac{\phi^2(F)}{F^2} \cdot (Y^T \Lambda^T Q \Lambda Y) \right] \\
&\leq \sigma^2 \text{ch}_1(\Lambda^T Q \Lambda V) E \left[\frac{\phi^2(F)}{F} - \frac{4}{(n_1 + n_2 - 2)p + 2} \phi(F)\phi'(F) \right] \tag{3.10}
\end{aligned}$$

Combining (3.3) and (3.10), one gets

$$\begin{aligned}
&\sigma^{-2} E[(\mu_1^\phi - \mu_1)^T Q(\mu_1^\phi - \mu_1) - (\bar{X}_1 - \mu_1)^T Q(\bar{X}_1 - \mu_1)] \\
&\leq -2E \left[\frac{\phi(F)}{F} \text{tr}(\Lambda^T Q \Lambda V) + 2 \left(\phi'(F) - \frac{\phi(F)}{F} \right) \frac{Y^T \Lambda^T Q \Lambda Y}{Y^T V^{-1} Y} \right] \\
&\quad + \text{ch}_1(\Lambda^T Q \Lambda V) E \left[\frac{\phi^2(F)}{F} - \frac{4}{(n_1 + n_2 - 2)p + 2} \phi(F)\phi'(F) \right] \tag{3.11}
\end{aligned}$$

The following theorem is now easy to prove from (3.11). Recall that $\Lambda = (n_1 V_1^{-1} + n_2 V_2^{-1})^{-1} n_2 V_2^{-1}$ and $V = n_1^{-1} V_1 + n_2^{-1} V_2$.

Theorem 3.2 Suppose that

- (i) $\text{tr}(\Lambda^T Q \Lambda V) > 2 \text{ch}_1(\Lambda^T Q \Lambda V)$
- (ii) $0 < \phi(F) < 2 \left[\frac{\text{tr}(\Lambda^T Q \Lambda V)}{\text{ch}_1(\Lambda^T Q \Lambda V)} - 2 \right]$ and
- (iii) $\phi(F) \uparrow$ in F

hold. Then $\sigma^{-2} E[(\mu_1^\phi - \mu_1)^T Q(\mu_1^\phi - \mu_1) - (\bar{X}_1 - \mu_1)^T Q(\bar{X}_1 - \mu_1)] < 0$ for all μ_1 and μ_2 .

Proof: Using (iii), it follows from (3.11) that

$$\begin{aligned}
 & \sigma^{-2} E[(\mu_1^\phi - \mu_1)^T Q(\mu_1^\phi - \mu_1) - (\bar{X}_1 - \mu_1)^T Q(\bar{X}_1 - \mu_1)] \\
 & \leq 2E \left[-\frac{\phi(F)}{F} \text{tr}(\Lambda^T Q \Lambda V) + 2 \frac{\phi(F)}{F} \cdot \frac{Y^T \Lambda^T Q \Lambda Y}{Y^T V^{-1} Y} + \frac{1}{2} \frac{\phi^2(F)}{F} \text{ch}_1(\Lambda^T Q \Lambda V) \right] \\
 & \leq 2E \left[-\frac{\phi(F)}{F} \text{tr}(\Lambda^T Q \Lambda V) + 2 \frac{\phi(F)}{F} \text{ch}_1(\Lambda^T Q \Lambda V) + \frac{1}{2} \frac{\phi^2(F)}{F} \text{ch}_1(l^T Q \Lambda V) \right] \\
 & = 2E \left[-\frac{\phi(F)}{2F} \text{ch}_1(\Lambda^T Q \Lambda V) \left\{ 2 \left(\frac{\text{tr}(\Lambda^T Q \Lambda V)}{\text{ch}_1(\Lambda^T Q \Lambda V)} - 2 \right) - \phi(F) \right\} \right] \\
 & < 0
 \end{aligned} \tag{3.12}$$

using conditions (i) and (ii) of the theorem.

Remark 3.1 It is an immediate consequence of the above theorem that if condition (i) of Theorem 3 holds, and $0 < p - 2 < 2[(\text{tr}(\Lambda^T Q \Lambda V)/\text{ch}_1(\Lambda^T Q \Lambda V) - 2)]$, then the EB estimators e_{EB} dominates \bar{X}_1 . In particular, if $Q = V_1 = V_2 = I_p$, then $\text{tr}(\Lambda^T Q \Lambda V) = p \text{ch}_1(\Lambda^T Q \Lambda V)$, and hence e_{EB} dominates \bar{X}_1 for $p \geq 3$.

In the remainder of this section we show how a modified EB estimator can dominate the PTE. Once again, an appeal to Theorem 3.1 is made.

A PTE δ_{PTE} of μ_1 is of the form $\delta_{PTE} = g(F)\bar{X}_1 + (1 - g(F))W = \bar{X}_1 - (1 - g(F))(\bar{X}_1 - W)$ where $g(F) = I_{[F > d]}$ for some positive constant d , and I denotes the usual indicator function. The choice of d is governed by the level of significance that is used for testing $H_0: \mu_1 = \mu_2$. We propose the rival estimator

$$\begin{aligned}
 \delta_{MEB} &= \bar{X}_1 - (1 - (1 - \frac{c}{F})g(F))(\bar{X}_1 - W) \\
 &= W + (1 - \frac{c}{F})g(F)(\bar{X}_1 - W)
 \end{aligned} \tag{3.13}$$

which is a modified version of e_{EB} with $p - 2$ replaced by a general c . Note that $\delta_{MEB} = W$ when $g(F) = 0$, but $\delta_{MEB} = \delta_{EB}$ when $g(F) = 1$. The following theorem is then proved.

Theorem 3.3 Suppose condition (i) of Theorem 3.2 holds and $0 < c < 2 \left[\frac{\text{tr}(\Lambda^T Q \Lambda V)}{\text{ch}_1(\Lambda^T Q \Lambda V)} - 2 \right]$. Then

$$\sigma^{-2} E[(\delta_{MEB} - \mu_1)^T Q(\delta_{MEB} - \mu_1) - (\delta_{PTE} - \mu_1)^T Q(\delta_{PTE} - \mu_1)] < 0 \tag{3.14}$$

for all μ_1 and μ_2 .

Proof:

Write $\phi_1(F) = F(1 - g(F))$ and $\phi_2(F) = F(1 - (1 - \frac{c}{F})g(F)) = \phi_1(F) + cg(F)$. Then $\delta_{PTE} = \bar{X}_1 - (\phi_1(F)/F)(\bar{X}_1 - W)$ while $\delta_{MEB} = \bar{X}_1 - (\phi_2(F)/F)(\bar{X}_1 - W)$. Note that both $\phi_1(F)$ and $\phi_2(F)$ are differentiable everywhere except at $F = d$. Thus $\phi'_1(F)$ and $\phi'_2(F)$ are defined a.e. (Lebesgue). Moreover, $\phi_1(F) - \phi_2(F) = -cg(F)$, $\phi_1^2(F) - \phi_2^2(F) = -c^2g^2(F) = -c^2g(F)$ and $\phi'_1(F) = \phi'_2(F) = 1 - g(F)$ a.e. Lebesgue. Then, applying Theorem 3.1 twice once with $\phi(F) = \phi_2(F)$, and next with $\phi(F) = \phi_1(F)$, one gets

Left hand side of (3.14)

$$\begin{aligned}
 &= -2E \left[\frac{cg(F)}{F} \text{tr}(\Lambda^T Q \Lambda V) - \frac{2}{F} cg(F) \frac{Y^T \Lambda^T Q \Lambda Y}{Y^T V^{-1} Y} \right] \\
 &+ \sigma^{-2} E \left[\frac{c^2 g^2(F)}{F^2} \cdot (Y^T \Lambda^T Q \Lambda Y) \right] \\
 &\leq -2E \left[\frac{cg(F)}{F} \text{tr}(\Lambda^T Q \Lambda V) - \frac{2}{F} cg(F) \text{ch}_1(\Lambda^T Q \Lambda V) \right] \\
 &+ \sigma^{-2} E \left[\frac{c^2 g^2(F)}{F^2} \cdot F \cdot \frac{\text{tr}(V_1^{-1} S_1 + V_2^{-1} S_2)}{(n_1 + n_2 - 2)p + 2} \cdot \frac{Y^T \Lambda^T Q \Lambda Y}{Y^T V^{-1} Y} \right] \\
 &\leq -2E \left[\frac{cg(F)}{F} \text{tr}(\Lambda^T Q \Lambda V) - \frac{2cg(F)}{F} \cdot \text{ch}_1(\Lambda^T Q \Lambda V) \right] \\
 &+ \sigma^{-2} \text{ch}_1(\Lambda^T Q \Lambda V) E \left[\frac{c^2 g^2(F)}{F^2} \cdot F \cdot \frac{\text{tr}(\Lambda^T Q \Lambda V)}{(n_1 + n_2 - 2)p + 2} \right] \quad (3.15)
 \end{aligned}$$

Applying (2.18) of Efron and Morris (1976) again with $\phi(F) = g(F)$ so that $\phi'(F) = 0$ a.e. Lebesgue, one gets

$$E \left[\left(\frac{g^2(F)}{F^2} \right) F \text{tr}(V_1^{-1} S_1 + V_2^{-1} S_2) / ((n_1 + n_2 - 2)p + 2) \right] = \sigma^2 E[g^2(F)/F] = \sigma^2 E[g(F)/F]. \quad (3.16)$$

Now from (3.15) and (3.16), left hand side of (3.14)

$$\begin{aligned}
 &\leq -E \left[\frac{cg(F)}{F} \text{ch}_1(\Lambda^T Q \Lambda V) \left\{ 2 \left(\frac{\text{tr}(\Lambda^T Q \Lambda V)}{\text{ch}_1(\Lambda^T Q \Lambda V)} - 2 \right) - c \right\} \right] \\
 &\leq 0 \quad (3.17)
 \end{aligned}$$

by using the upper bound of c given in this theorem. The proof of the theorem is complete.

Remark 3.2 *Note that when $Q = V_1 = V_2 = I_p$, the conditions of the theorem hold when $0 < c < 2(p - 2)$, and in particular when $c = p - 2$, $p \geq 3$.*

4. Hierarchical Bayes Estimation

Section 2 is devoted to classical empirical Bayes estimation, i.e. when the unknown prior parameters are estimated by classical methods of estimation such as uniformly minimum variance unbiased estimation, maximum likelihood estimation, best invariant estimation etc. Instead, one can assign prior distributions (proper or improper) to the hyperparameters, and come up with hierarchical Bayes (HB) estimators of μ_1 . Note that in a classical EB approach, the lower stage Bayesian analysis is performed as if the hyperparameters were known a priori. This approach ignores the error associated with the estimation of the hyperparameters. On the other hand, the HB approach models the uncertainty of the hyperparameters by the second stage prior. Accordingly, unlike positive part EB estimators, the HB estimators are smooth, and bear the potentiality of being admissible.

To introduce the HB model, first note that as in Section 2, one may start with the minimal sufficient statistic $(\bar{X}_1, \bar{X}_2, \text{tr}(V_1^{-1}S_1 + V_2^{-1}S_2))$. Write $r^{-1} = \sigma^2$ and $(\rho r)^{-1} = \tau^2$ i.e. $\rho = \sigma^2/\tau^2$. Now conditional on μ_1, μ_2 and r , \bar{X}_1, \bar{X}_2 and $U = \text{tr}(V_1^{-1}S_1 + V_2^{-1}S_2)$ are mutually independent with $\bar{X}_1 \sim N_p(\mu_1, (n_1 r)^{-1}V_1)$, $\bar{X}_2 \sim N_p(\mu_2, (n_2 r)^{-1}V_2)$ and $U \sim r^{-1}\chi_{(n_1+n_2-2)p}^2$. Next we assume that conditional on ν, ρ and r , μ_1 and μ_2 are mutually independent with $\mu_1 \sim N(\nu, (r\rho)^{-1}n_1^{-1}V_1)$ and $\mu_2 \sim N(\nu, n_2^{-1}(\rho r)^{-1}V_2)$. Also, it is assumed that ν, ρ and r are mutually independent with ν uniform on R^p , ρ has the type II Beta distribution with pdf $h_1(\rho) \propto \rho^{m-1}(1+\rho)^{-(m+1)}I_{[\rho>0]}$ where $m(>0)$ is known, while r has a gamma distribution with pdf $h_2(r) \propto \exp(-\frac{1}{2}\alpha r)r^{\delta-1}$, $\alpha(>0)$ and $\delta(>0)$ being known. We shall aim at finding the posterior distribution of $\mu = (\mu_1^T, \mu_2^T)^T$ given \bar{X}_1, \bar{X}_2 and u .

First note that the joint prior distribution of μ_1, μ_2, ν, r and ρ is given by

$$\begin{aligned} f(\mu_1, \mu_2, \nu, r, \rho) &\propto (\rho r)^p \\ &\cdot \exp\left[-\frac{\rho r}{2}\{n_1(\mu_1 - \nu)^T V_1^{-1}(\mu_1 - \nu) + n_2(\mu_2 - \nu)^T V_2^{-1}(\mu_2 - \nu)\}\right] \\ &\cdot h_1(\rho)h_2(r) \end{aligned} \quad (4.1)$$

Next observe that

$$\begin{aligned} n_1(\mu_1 - \nu)^T V_1^{-1}(\mu_1 - \nu) + n_2(\mu_2 - \nu)^T V_2^{-1}(\mu_2 - \nu) \\ = [(\nu - \mu_*)^T V_*^{-1}(\nu - \mu_*)] \\ + n_1\mu_1^T V_1^{-1}\mu_1 + n_2\mu_2^T V_2^{-1}\mu_2 - \mu_*^T V_*^{-1}\mu_*, \end{aligned} \quad (4.2)$$

where one may recall that $\mu_* = (n_1 V_1^{-1} + n_2 V_2^{-1})^{-1}(n_1 V_1^{-1}\mu_1 + n_2 V_2^{-1}\mu_2) = (V_*^{-1})^{-1}(n_1 V_1^{-1}\mu_1 + n_2 V_2^{-1}\mu_2)$ with $V_*^{-1} = n_1 V_1^{-1} + n_2 V_2^{-1}$. Now integrating with respect to ν , one gets the joint pdf of μ_1, μ_2, r and ρ in the form $f(\mu_1, \mu_2, r, \rho)$

$$\propto (\rho r)^{p/2} \exp\left[-\frac{\rho r}{2}\{n_1\mu_1^T V_1^{-1}\mu_1 + n_2\mu_2^T V_2^{-1}\mu_2 - \mu_*^T V_*^{-1}\mu_*\}\right] h_1(\rho)h_2(r) \quad (4.3)$$

The exponent in (4.3) is easily simplified as

$$\begin{aligned} & n_1 \mu_1^T V_1^{-1} \mu_1 + n_2 \mu_2^T V_2^{-1} \mu_2 - \mu^T V_*^{-1} \mu, \\ &= \mu_1^T \{n_1 V_1^{-1} - n_1 V_1^{-1} V_* n_1 V_1^{-1}\} \mu_1 + \mu_2^T \{n_2 V_2^{-1} - n_2 V_2^{-1} V_* n_2 V_2^{-1}\} \\ &= \mu_1^T n_1 V_1^{-1} V_* n_2 V_2^{-1} \mu_2 - \mu_2^T n_2 V_2^{-1} V_* n_1 V_1^{-1} \mu_1 \end{aligned} \quad (4.4)$$

where $V_* = (n_1 V_1^{-1} + n_2 V_2^{-1})^{-1}$. Also, the joint pdf of \bar{X}_1, \bar{X}_2 and U conditional on μ_1, μ_2 and r is given by

$$\begin{aligned} & f(\bar{x}_1, \bar{x}_2, u | \mu_1, \mu_2, r) \\ & \propto r^p \exp[-r/2 \{n_1(\bar{x}_1 - \mu_1)^T V_1^{-1}(\bar{x}_1 - \mu_1) + n_2(\bar{x}_2 - \mu_2)^T V_2^{-1}(\bar{x}_2 - \mu_2)\}] \\ & \quad \cdot \exp(-ru/2) u^{(n_1+n_2-2)p/2-1} r^{(n_1+n_2-2)p/2} \end{aligned} \quad (4.5)$$

Next we calculate

$$\begin{aligned} G &= n_1(\mu_1 - \bar{x}_1)^T V_1^{-1}(\mu_1 - \bar{x}_1) + n_2(\mu_2 - \bar{x}_2)^T V_2^{-1}(\mu_2 - \bar{x}_2) \\ &+ \rho \{n_1 \mu_1^T V_1^{-1} \mu_1 + n_2 \mu_2^T V_2^{-1} \mu_2 - \mu^T V_*^{-1} \mu\} \end{aligned} \quad (4.6)$$

which is needed to derive the posterior distribution of μ given \bar{x}_1, \bar{x}_2 and u . Using (4.4) and straightforward algebra, one gets

$$\begin{aligned} G &= \mu_1^T D_{11} \mu_1 + \mu_2^T D_{22} \mu_2 - 2\mu_1^T D_{12} \mu_2 - 2n_1 \bar{x}_1^T V_1^{-1} \mu_1 - 2n_2 \bar{x}_2^T V_2^{-1} \mu_2 \\ &+ n_1 \bar{x}_1^T V_1^{-1} \bar{x}_1 + n_2 \bar{x}_2^T V_2^{-1} \bar{x}_2 \end{aligned} \quad (4.7)$$

where

$$\begin{aligned} D_{11} &= n_1 V_1^{-1} + \rho \{n_1 V_1^{-1} - n_1 V_1^{-1} V_* n_1 V_1^{-1}\}, \\ D_{22} &= n_2 V_2^{-1} + \rho \{n_2 V_2^{-1} - n_2 V_2^{-1} V_* n_2 V_2^{-1}\}, \\ D_{12} &= \rho n_1 V_1^{-1} V_* n_2 V_2^{-1} \end{aligned} \quad (4.8)$$

We now write G as $G_1 + G_2$ where

$$\begin{aligned} G_1 &= [(\mu_1 - A_{11} \bar{x}_1 - A_{12} \bar{x}_2)^T D_{11} (\mu_1 - A_{11} \bar{x}_1 - A_{12} \bar{x}_2) \\ &+ (\mu_2 - A_{21} \bar{x}_1 - A_{22} \bar{x}_2)^T D_{22} (\mu_2 - A_{21} \bar{x}_1 - A_{22} \bar{x}_2) \\ &- 2(\mu_1 - A_{11} \bar{x}_1 - A_{12} \bar{x}_2)^T D_{12} (\mu_2 - A_{21} \bar{x}_1 - A_{22} \bar{x}_2)] \end{aligned} \quad (4.9)$$

and

$$\begin{aligned}
G_2 = & [n_1 \bar{\mathbf{x}}_1^T V_1^{-1} \bar{\mathbf{x}}_1 + n_2 \bar{\mathbf{x}}_2^T V_2^{-1} \bar{\mathbf{x}}_2 - (A_{11} \bar{\mathbf{x}}_1 + A_{12} \bar{\mathbf{x}}_2)^T D_{11} (A_{11} \bar{\mathbf{x}}_1 + A_{12} \bar{\mathbf{x}}_2) \\
& - (A_{21} \bar{\mathbf{x}}_1 + A_{22} \bar{\mathbf{x}}_2)^T D_{22} (A_{21} \bar{\mathbf{x}}_1 + A_{22} \bar{\mathbf{x}}_2) \\
& + 2(A_{11} \bar{\mathbf{x}}_1 + A_{12} \bar{\mathbf{x}}_2)^T D_{12} (A_{21} \bar{\mathbf{x}}_1 + A_{22} \bar{\mathbf{x}}_2)]
\end{aligned} \tag{4.10}$$

From (4.7), (4.9) and (4.10), it follows that A_{11}, A_{12}, A_{21} and A_{22} satisfy

$$\begin{aligned}
D_{11} A_{11} &= D_{12} A_{21} = n_1 V_1^{-1} \\
D_{22} A_{22} &= D_{12}^T A_{12} = n_2 V_2^{-1} \\
D_{11} A_{12} &= D_{12} A_{22} \\
D_{22} A_{21} &= D_{12}^T A_{11}
\end{aligned} \tag{4.11}$$

which can be rewritten as

$$\begin{aligned}
A_{12} - D_{11}^{-1} D_{12} A_{22}, A_{21} &= D_{22}^{-1} D_{12}^T A_{11}, \\
(D_{11} - D_{12} D_{22}^{-1} D_{12}^T) A_{11} &= n_1 V_1^{-1}, \\
(D_{22} - D_{12}^T D_{11}^{-1} D_{12}) A_{22} &= n_2 V_2^{-1}
\end{aligned} \tag{4.12}$$

The following lemma whose proof is given in the Appendix is crucial to further simplify G_2 . Recall that $B = \sigma^2/(\sigma^2 + \tau^2) = \rho/(1 + \rho)$ and $\mathbf{W} = (n_1 V_1^{-1} + n_2 V_2^{-1})^{-1} (n_1 V_1^{-1} \bar{\mathbf{x}}_1 + n_2 V_2^{-1} \bar{\mathbf{x}}_2) = V_1 (n_1 V_1^{-1} \bar{\mathbf{x}}_1 + n_2 V_2^{-1} \bar{\mathbf{x}}_2)$.

Lemma 4.1

$$A_{11} \bar{\mathbf{x}}_1 + A_{12} \bar{\mathbf{x}}_2 = (1 - B) \bar{\mathbf{x}}_1 + B \mathbf{W} =: \mathbf{h}_1(\text{say}) \tag{4.13}$$

$$A_{21} \bar{\mathbf{x}}_1 + A_{22} \bar{\mathbf{x}}_2 = (1 - B) \bar{\mathbf{x}}_2 + B \mathbf{W} =: \mathbf{h}_2(\text{say}) \tag{4.14}$$

From (4.10), (4.13) and (4.14), G_2 can be simplified as

$$\begin{aligned}
G_2 = & n_1 \bar{\mathbf{x}}_1^T V_1^{-1} \bar{\mathbf{x}}_1 + n_2 \bar{\mathbf{x}}_2^T V_2^{-1} \bar{\mathbf{x}}_2 \\
& \{ (1 - B) \bar{\mathbf{x}}_1^T + B \mathbf{W}^T \} D_{11} \{ (1 - B) \bar{\mathbf{x}}_1 + B \mathbf{W} \} \\
& \{ (1 - B) \bar{\mathbf{x}}_2^T + B \mathbf{W}^T \} D_{22} \{ (1 - B) \bar{\mathbf{x}}_2 + B \mathbf{W} \} \\
& + 2 \{ (1 - B) \bar{\mathbf{x}}_1^T + B \mathbf{W}^T \} D_{12} \{ (1 - B) \bar{\mathbf{x}}_2 + B \mathbf{W} \} \\
= & \bar{\mathbf{x}}_1^T [n_1 V_1^{-1} - \{ (1 - B) I \\
& + B n_1 V_1^{-1} V_1 \} D_{11} \{ (1 - B) I + B n_1 V_1 V_1^{-1} \} \\
& (B n_1 V_1^{-1} V_1) D_{22} (B n_1 V_1 V_1^{-1})
\end{aligned}$$

$$\begin{aligned}
& + 2\{(1-B)I + Bn_1V_1^{-1}V_*\}D_{12}(Bn_1V_*V_1^{-1})|\bar{\mathbf{x}}_1 \\
& + \bar{\mathbf{x}}_2^T[n_2V_2^{-1} - \{(1-B)I \\
& + Bn_2V_2^{-1}V_*\}D_{22}\{(1-B)I + Bn_2V_*V_2^{-1}\} \\
& - (Bn_2V_2^{-1}V_*)D_{11}(Bn_2V_*V_2^{-1}) + 2\{(1-B)I \\
& + Bn_2V_2^{-1}V_*\}D_{12}^T(Bn_2V_*V_2^{-1})|\bar{\mathbf{x}}_2 \\
& - \bar{\mathbf{x}}_1^T[\{(1-B)I + Bn_1V_1^{-1}V_*\}D_{11}(Bn_2V_*V_2^{-1}) \\
& + (Bn_1V_1^{-1}V_*)D_{22}(1-B)I \\
& + Bn_2V_*V_2^{-1}\} + 2\{(1-B)I + Bn_1V_1^{-1}V_*\}D_{12}\{(1-B)I \\
& + Bn_2V_*V_2^{-1}\}|\bar{\mathbf{x}}_2 \\
& - \bar{\mathbf{x}}_2^T[(Bn_2V_2^{-1}V_*)D_{11}\{(1-B)I \\
& + Bn_1V_*V_1^{-1}\} + \{(1-B)I + Bn_2V_2^{-1}V_*\}D_{22} \\
& + B(n_1V_*V_1^{-1}) + 2(Bn_2V_2^{-1}V_*)D_{12}(Bn_1V_*V_1^{-1})|\bar{\mathbf{x}}_1
\end{aligned} \tag{4.15}$$

From (4.8), one gets

$$\begin{aligned}
D_{11} + D_{22} - 2D_{12} &= (1+\rho)(n_1V_1^{-1} + n_2V_2^{-1}) - \rho(n_1V_1^{-1} + n_2V_2^{-1})V_*(n_1V_1^{-1} + n_2V_2^{-1}) \\
&= n_1V_1^{-1} + n_2V_2^{-1} \text{ (since } V_*^{-1} = n_1V_1^{-1} + n_2V_2^{-1}) \\
&= V_*^{-1}
\end{aligned} \tag{4.16}$$

Using (4.8) and (4.16), it is possible to simplify G_2 considerably. This is done in the following Lemma whose proof appears in the Appendix.

Lemma 4.2 $G_2 = B[\bar{\mathbf{x}}_1^T\{n_1V_1^{-1} - n_1V_1^{-1}V_*n_1V_1^{-1}\}\bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2^T\{n_2V_2^{-1} - n_2V_2^{-1}V_*n_2V_2^{-1}\}\bar{\mathbf{x}}_2 + 2\bar{\mathbf{x}}_1^T(n_1V_1^{-1}V_*n_2V_2^{-1})\bar{\mathbf{x}}_2]$

Therefore, from (4.9), Lemma 4.1 and Lemma 4.2, G can be written as

$$\begin{aligned}
G &= (\mu_1 - \mathbf{b}_1)^T D_{11}(\mu_1 - \mathbf{b}_1) + (\mu_2 - \mathbf{b}_2)^T D_{22}(\mu_2 - \mathbf{b}_2) - 2(\mu_1 - \mathbf{b}_1)^T D_{12}(\mu_2 - \mathbf{b}_2) \\
&+ B[\bar{\mathbf{x}}_1^T D_{11}^* \bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2^T D_{22}^* \bar{\mathbf{x}}_2 + 2\bar{\mathbf{x}}_1^T D_{12}^* \bar{\mathbf{x}}_2]
\end{aligned} \tag{4.17}$$

where

$$\begin{aligned}
D_{11}^* &= n_1V_1^{-1} - n_1V_1^{-1}V_*n_1V_1^{-1} \\
D_{22}^* &= n_2V_2^{-1} - n_2V_2^{-1}V_*n_2V_2^{-1} \\
D_{12}^* &= n_1V_1^{-1}V_*n_2V_2^{-1}
\end{aligned} \tag{4.18}$$

Returning to (4.3) and (4.5), the joint pdf of $\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, u, \mu_1, \mu_2, r$ and ρ is given by

$$\begin{aligned}
& f(\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, u, \mu_1, \mu_2, r, \rho) \\
& \propto r^r (\rho r)^{r/2} \cdot \exp\left[-\frac{r}{2}G\right] \cdot \exp[-ru/2] \\
& u^{(n_1+n_2-2)r/2-1} r^{(n_1+n_2-2)r/2} h_1(\rho) h_2(r)
\end{aligned} \tag{4.19}$$

It follows from (4.17) and (4.19) that conditional on $\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, u, r$ and ρ ,

$$\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \sim N_{2p} \left[\begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix}, r^{-1} \begin{pmatrix} D_{11} & -D_{12} \\ -D_{12}^T & D_{22} \end{pmatrix}^{-1} \right] \tag{4.20}$$

Also, integrating out with respect to μ_1 and μ_2 , it follows from (4.19) that the joint pdf of $\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2, U, r$ and ρ is given by

$$\begin{aligned}
f(\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, u, r, \rho) & \propto (\rho r)^{p/2} \left| \begin{pmatrix} D_{11} & -D_{12} \\ -D_{12}^T & D_{22} \end{pmatrix} \right|^{-\frac{1}{2}} \exp\left[-\frac{r}{2}\{u + BSS_H\}\right] \\
& \cdot r^{(n_1+n_2-2)r/2} \cdot u^{(n_1+n_2-2)r/2-1} \cdot \rho^{m-1} (1+\rho)^{-(m+1)} \cdot \exp(-\alpha r/2) r^{\delta-1}
\end{aligned} \tag{4.21}$$

where

$$SS_H = \bar{\mathbf{x}}_1^T D_{11} + \bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2^T D_{22} + \bar{\mathbf{x}}_2 - 2\bar{\mathbf{x}}_1^T D_{12} + \bar{\mathbf{x}}_2 \tag{4.22}$$

Now, from (4.8), one gets

$$\begin{aligned}
& \left| \begin{pmatrix} D_{11} & -D_{12} \\ D_{12}^T & D_{22} \end{pmatrix} \right| \\
& = \left| (1+\rho) \begin{pmatrix} n_1 V_1^{-1} & 0 \\ 0 & n_2 V_2^{-1} \end{pmatrix} - \rho \begin{pmatrix} n_1 V_1^{-1} V_1 n_1 V_1^{-1} & n_1 V_1^{-1} V_1 n_2 V_2^{-1} \\ n_2 V_2^{-1} V_1 n_1 V_1^{-1} & n_2 V_2^{-1} V_1 n_2 V_2^{-1} \end{pmatrix} \right| \\
& = (1+\rho)^{2r} \left| \begin{pmatrix} n_1 V_1^{-1} & 0 \\ 0 & n_2 V_2^{-1} \end{pmatrix} - B \begin{pmatrix} n_1 V_1^{-1} \\ n_2 V_2^{-1} \end{pmatrix} V_1 (n_1 V_1^{-1} : n_2 V_2^{-1}) \right| \\
& = (1+\rho^{2r}) \left| \begin{pmatrix} n_1 V_1^{-1} & 0 \\ 0 & n_2 V_2^{-1} \end{pmatrix} \right| \\
& \cdot \left| I_{2r} - B \begin{pmatrix} \frac{V_1^{-1}}{n_1} & 0 \\ 0 & \frac{V_2^{-1}}{n_2} \end{pmatrix} \begin{pmatrix} n_1 V_1^{-1} \\ n_2 V_2^{-1} \end{pmatrix} V_1 (n_1 V_1^{-1} : n_2 V_2^{-1}) \begin{pmatrix} \frac{V_1^{-1}}{n_1} & 0 \\ 0 & \frac{V_2^{-1}}{n_2} \end{pmatrix} \right| \\
& = (1+\rho)^{2r} |n_1 V_1^{-1}| |n_2 V_2^{-1}| \left| I_r - B(n_1 V_1^{-1} n_2 V_2^{-1}) \begin{pmatrix} \frac{V_1}{n_1} & 0 \\ 0 & \frac{V_2}{n_2} \end{pmatrix} \begin{pmatrix} n_1 V_1^{-1} \\ n_2 V_2^{-1} \end{pmatrix} V_1 \right| \\
& = (1+\rho)^{2r} |n_1 V_1^{-1}| |n_2 V_2^{-1}| \left| I_r - B I_r \right| \quad (\text{since } V_1^{-1} = n_1 V_1^{-1} + n_2 V_2^{-1}) \\
& = (1+\rho)^r |n_1 V_1^{-1}| |n_2 V_2^{-1}| \tag{4.23}
\end{aligned}$$

Hence, from (4.21) and (4.23), one gets

$$f(\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, u, r, \rho) \propto \rho^{p/2} r^{(n_1+n_2-1)p/2} (1+\rho)^{-p/2} \exp\left[-\frac{r}{2}(u + BSS_H + \alpha)\right] \\ \cdot u^{(n_1+n_2-2)p/2-1} \cdot r^{\delta-1} \rho^{m-1} (1+\rho)^{-(m+1)} \quad (4.24)$$

Integrating out with respect to r , one gets the joint pdf of $\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2, U$ and ρ as

$$f(\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, u, \rho) \propto \left(\frac{\rho}{1+\rho}\right)^{p/2} u^{(n_1+n_2-2)p/2-1} (u + BSS_H + \alpha)^{-(n_1+n_2-1)p/2-\delta} \\ \cdot \rho^{m-1} (1+\rho)^{-(m+1)} \quad (4.25)$$

Using the transformation $\rho/(1+\rho) = B$ provides the joint pdf of $\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2, U$ and B as

$$f(\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, u, B) \propto B^{p/2+m-1} u^{(n_1+n_2-2)p/2-1} (u + BSS_H + \alpha)^{-(n_1+n_2-1)p/2-\delta} \quad (4.26)$$

Next observe from (4.20) and (4.13) that

$$E(\mu_1 | B, \bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, u, r) = \mathbf{b}_1 = (1-B)\bar{\mathbf{x}}_1 + B\mathbf{W}.$$

Hence the HJB estimator of μ_1 is

$$E(\mu_1 | \bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, u) = \bar{\mathbf{x}}_1 + E(B | \bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, u)(\bar{\mathbf{x}}_1 - \mathbf{W}) \quad (4.27)$$

But, from (4.26), one gets

$$E(B | \bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, u) = \frac{\int_0^1 B^{p/2+m} (u + BSS_H + \alpha)^{-(n_1+n_2-1)p/2-\delta} dB}{\int_0^1 B^{p/2+m-1} (u + BSS_H + \alpha)^{-(n_1+n_2-1)p/2-\delta} dB} \quad (4.28)$$

Remark 4.1 From simultaneous diagonalization of $n_1 V_1^{-1}$ and $n_2 V_2^{-1}$, it is easy to show from (4.18) that

$$D_{11}^* = D_{22}^* = D_{12}^* = (n_1 V_1^{-1} + n_2 V_2^{-1})^{-1}, \quad (4.29)$$

so that from (4.22) one gets

$$SS_H = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)^T (n_1 V_1^{-1} + n_2 V_2^{-1})^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) \quad (4.30)$$

which is precisely the numerator of F defined in (3.1).

Remark 4.2 It is sometimes possible to reduce the above IIB estimator to an EB estimator of the form $\bar{\mathbf{x}}_1 - (\phi(F)/F)(\bar{\mathbf{x}}_1 - \mathbf{W})$. Consider for example the situation when $\alpha = 0$ i.e., R has the improper prior $h_2(r) = r^{\delta-1}$. Now writing $v = SS_H/u$, we note from (4.30) that $F = ((n_1 + n_2 - 2)p + 2)v$. Also, for $\alpha = 0$, it follows from (4.28) that

$$\begin{aligned} E(B|\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, u) &= \int_0^1 B^{p/2+m}(1+Bv)^{-(n_1+n_2-2)p-\delta} dB / \int_0^1 B^{p/2+m-1}(1+Bv)^{-(n_1+n_2-2)p-\delta} dB \\ &= v^{-1} \int_0^1 \left(\frac{1}{1+Bv} \right)^{(n_1+n_2-2)p+\delta-m-2} \left(\frac{Bv}{1+Bv} \right)^{p/2+m} \frac{v dB}{(1+Bv)^2} \\ &\quad + \int_0^1 \left(\frac{Bv}{1+Bv} \right)^{p/2+m-1} \left(\frac{1}{1+Bv} \right)^{(n_1+n_2-2)p+\delta-m-1} \frac{v dB}{(1+Bv)^2} \\ &= v^{-1} \int_0^{v/(1+v)} u^{p/2+m} (1-u)^{(n_1+n_2-2)p+\delta-m-2} du \\ &\quad + \int_0^{v/(1+v)} u^{p/2+m-1} (1-u)^{(n_1+n_2-2)p+\delta-m-1} du \end{aligned} \quad (4.31)$$

From (4.31) it follows that $E(B|\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, u)$ can be expressed as $\phi(v)/v = \phi(F)/F$. Next note that integration by parts gives numerator of (4.31)

$$\begin{aligned} &= v^{-1} \left\{ \left(\frac{v}{1+v} \right)^{p/2+m} (1+v)^{-(n_1+n_2-2)p+\delta+m+1} \right. \\ &\quad \left. + \frac{\{ \frac{p}{2} + m \}}{v \{ (n_1+n_2-2)p+\delta-m-1 \}} \int_0^{v/(1+v)} u^{p/2+m-1} (1-u)^{(n_1+n_2-2)p+\delta-m-1} du \right. \\ &\quad \left. - \frac{p+2m}{v \{ (n_1+n_2-2)p+2\delta-2m-2 \}} \int_0^{v/(1+v)} u^{p/2+m-1} (1-u)^{(n_1+n_2-2)p+\delta-m-1} du \right\} \end{aligned} \quad (4.32)$$

Hence from (4.31) and (4.32),

$$\begin{aligned} E(B|\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, u) &= \frac{p+2m}{v \{ (n_1+n_2-2)p+2\delta-2m-2 \}} \\ &\quad - \frac{(p+2m)((n_1+n_2-2)p+2)}{F \{ (n_1+n_2-2)p+2\delta-2m-2 \}} \end{aligned} \quad (4.33)$$

so that

$$\begin{aligned} \phi(F) &= \frac{(p+2m)((n_1+n_2-2)p+2)}{(n_1+n_2-2)p+2\delta-2m-2} \\ &\quad - 2(p-2) \end{aligned}$$

if $(p + 2m)((n_1 + n_2 - 2)p + 2) < 2(p - 2)((n_1 + n_2 - 2)p - 2m - 2)$ ($\because \delta > c$) \iff $p\{2m(n_1 + n_2) + 6\} < p(p - 4)(n_1 + n_2 - 2) + 4m + 8$ which holds whenever $p \geq 5$ and $m < \{(p - 4)(n_1 + n_2 - 2) - 6\}/2(n_1 + n_2)$, assuming $n_1 + n_2 \geq 8$. Hence, for this choice of m , $\phi(F)$ satisfies condition (ii) of Theorem 3.2 for $Q = V_1 = V_2 = I_p$. Also, for $Q = V_1 = V_2 = I_p$, condition (i) of Theorem 3.2 automatically holds when $p \geq 3$.

Finally, noting that v is strictly increasing in F , and using the inequality

$$\begin{aligned} & \int_0^{v/(1+v)} u^{\frac{p}{2}+m} (1-u)^{-(n_1+n_2-2)p+\delta-m-2} du \\ &= \int_0^{v/(1+v)} \binom{u}{1-u} u^{\frac{p}{2}+m-1} (1-u)^{-(n_1+n_2-2)p+\delta-m-1} du \\ &= v \int_0^{v/(1+v)} u^{\frac{p}{2}+m-1} (1-u)^{-(n_1+n_2-2)p+\delta-m-1} du, \end{aligned} \quad (4.34)$$

one gets after direct differentiation $\phi^*(v)' \geq 0$. Hence $\phi^*(v)$ is \uparrow in v . Hence, condition (iii) of Theorem 3.2 also holds. Therefore, when $\alpha = 0$, $Q = V_1 = V_2 = I_p$, $p \geq 5$ and $0 < m < \{(p - 4)(n_1 + n_2 - 2) - 6\}/2(n_1 + n_2)$, the HB estimator obtained in (4.27) is minimax.

Remark 4.3 The conclusion given in Remark 4.2 bears strong resemblance to Strawderman (1971) in the one sample problem. However, the formulation here is much more general than the one given in Strawderman (1971) or Strawderman (1973). First, the estimator is not shrunk towards zero or a prespecified point, but is shrunk towards the pooled mean. In Strawderman (1971), r is assumed to be known, whereas in Strawderman (1973), r is assumed to belong to (γ, ∞) for some $\gamma > 0$. Our formulation is also more general than the one given in Morris (1983) because there r is assumed known and $(pr)^{-1}$ is given a uniform prior on $(0, \infty)$.

A. Proofs

Proof of Lemma 4.1: To prove (4.13), note from (4.8) that

$$\begin{aligned}
 (D_{22} - D_{12}^T D_{11}^{-1} D_{12}) D_{12}^{-1} D_{11} &= D_{22} D_{12}^{-1} D_{11} - D_{12}^T \\
 &= n_2 V_2^{-1} (1 + \rho) \{I - BV, n_2 V_2^{-1}\} \cdot \frac{1}{\rho} \cdot \frac{V_2}{n_2} \cdot V_*^{-1} \frac{V_1}{n_1} \\
 &\quad \cdot n_1 V_1^{-1} (1 + \rho) \{I - BV, n_1 V_1^{-1}\} - \rho n_2 V_2^{-1} V, n_1 V_1^{-1} \\
 &= ((1 + \rho)^2 / \rho) (I - B n_2 V_2^{-1} V, V_*^{-1} (I - BV, n_1 V_1^{-1}) - \rho n_2 V_2^{-1} V, n_1 V_1^{-1} \\
 &= ((1 + \rho)^2 / \rho) (I - B n_2 V_2^{-1} V, ((1 - B) n_1 V_1^{-1} + n_2 V_2^{-1}) - \rho n_2 V_2^{-1} V, n_1 V_1^{-1} \\
 &= ((1 + \rho)^2 / \rho) \{((1 - B) n_1 V_1^{-1} + n_2 V_2^{-1} - B(1 - B) n_2 V_2^{-1} V, n_1 V_1^{-1} \\
 &\quad - B n_2 V_2^{-1} V, n_2 V_2^{-1}\} - \rho n_2 V_2^{-1} V, n_1 V_1^{-1} \\
 &= ((1 + \rho)^2 / \rho) \{(1 - B) n_1 V_1^{-1} + n_2 V_2^{-1} - B n_2 V_2^{-1} + B^2 n_2 V_2^{-1} V, n_1 V_1^{-1}\} - \\
 &\quad \rho n_2 V_2^{-1} V, n_1 V_1^{-1} \\
 &= (n_1 V_1^{-1} + n_2 V_2^{-1}) / B
 \end{aligned} \tag{A.1}$$

so that

$$D_{11}^{-1} D_{12} (D_{22} - D_{12}^T D_{11}^{-1} D_{12})^{-1} = BV, \tag{A.2}$$

and

$$D_{22} - D_{12}^T D_{11}^{-1} D_{12} = (n_1 V_1^{-1} + n_2 V_2^{-1}) D_{11}^{-1} D_{12} / B$$

which yields by symmetry

$$D_{11} - D_{12} D_{22}^{-1} D_{12}^T = (n_1 V_1^{-1} + n_2 V_2^{-1}) D_{22}^{-1} D_{12}^T / B \tag{A.3}$$

Hence from (4.8) one gets

$$\begin{aligned}
 (D_{11} - D_{12} D_{22}^{-1} D_{12}^T)^{-1} &= B (D_{12}^T)^{-1} D_{22} (n_1 V_1^{-1} + n_2 V_2^{-1})^{-1} \\
 &= B \frac{V_1}{n_1} V_*^{-1} \frac{V_2}{n_2} \cdot \frac{1}{\rho} \cdot (1 + \rho) n_2 V_2^{-1} (I - BV, n_2 V_2^{-1}) V_* \\
 &= \frac{V_1}{n_1} V_*^{-1} (I - BV, n_2 V_2^{-1}) V_* \\
 &= \frac{V_1}{n_1} - B \frac{V_1}{n_1} n_2 V_2^{-1} V,
 \end{aligned} \tag{A.4}$$

so that

$$(D_{11} - D_{12} D_{22}^{-1} D_{12}^T)^{-1} n_1 V_1^{-1} = I - B \frac{V_1}{n_1} n_2 V_2^{-1} V, n_1 V_1^{-1} \tag{A.5}$$

which yields from (A.2)

$$\begin{aligned}
 (D_{11} - D_{12}D_{22}^{-1}D_{12}^T)^{-1}n_1V_1^{-1} &= D_{11}^{-1}D_{12}(D_{22} - D_{12}^TD_{11}^{-1}D_{12})^{-1}n_1V_1^{-1} \\
 &= I - B\frac{V_1}{n_1}n_2V_2^{-1}V,n_1V_1^{-1} - BV,n_1V_1^{-1} \\
 &= I - B\frac{V_1}{n_1}(n_2V_2^{-1} + n_1V_1^{-1})V,n_1V_1^{-1} \\
 &= (1 - B)I
 \end{aligned} \tag{A.6}$$

(4.13) now follows upon writing from (4.12)

$$\begin{aligned}
 A_{11}\bar{x}_1 + A_{12}\bar{x}_2 &= (D_{11} - D_{12}D_{22}^{-1}D_{12}^T)^{-1}n_1V_1^{-1}\bar{x}_1 + D_{11}^{-1}D_{12}(D_{22} - D_{12}^TD_{11}^{-1}D_{12})^{-1}n_2V_2^{-1}\bar{x}_2 \\
 &= [(D_{11} - D_{12}D_{22}^{-1}D_{12}^T)^{-1}n_1V_1^{-1} - D_{11}^{-1}D_{12}(D_{22} - D_{12}^TD_{11}^{-1}D_{12})^{-1}n_1V_1^{-1}]\bar{x}_1 \\
 &\quad + D_{11}^{-1}D_{12}(D_{22} - D_{12}^TD_{11}^{-1}D_{12})^{-1}(n_2V_2^{-1}\bar{x}_2 + n_1V_1^{-1}\bar{x}_1)
 \end{aligned}$$

and invoking (A.2) and (A.6). The proof of (4.14) is similar and hence omitted.

Proof of Lemma 4.2: The matrix appearing in $\bar{x}_1^T[\cdot]\bar{x}_1$ in (4.15) is simplified as

$$\begin{aligned}
 n_1V_1^{-1} &\{ (1 - B)I + Bn_1V_1^{-1}V, \} D_{11} \{ (1 - B)I + Bn_1V,V_1^{-1} \} \\
 &\quad (Bn_1V_1^{-1}V,)D_{22}(Bn_1V,V_1^{-1}) + 2\{ (1 - B)I + Bn_1V_1^{-1}V, \} D_{12}(Bn_1V,V_1^{-1}) \\
 &= n_1V_1^{-1} - (1 - B)^2D_{11} - B(1 - B)(n_1V_1^{-1}V,D_{11} + D_{11}n_1V,V_1^{-1}) \\
 &\quad + 2B(1 - B)D_{12}n_1V,V_1^{-1} \\
 &\quad \quad B^2n_1V_1^{-1}V,n_1V_1^{-1} \\
 &= n_1V_1^{-1} - (1 - B)^2\{ n_1V_1^{-1}(1 + \rho) \\
 &\quad \quad \rho n_1V_1^{-1}V,n_1V_1^{-1} \} \\
 &\quad B(1 - B)n_1V_1^{-1}V,\{ n_1V_1^{-1}(1 + \rho) \\
 &\quad \quad \rho n_1V_1^{-1}V,n_1V_1^{-1} \} \\
 &\quad B(1 - B)\{ n_1V_1^{-1}(1 + \rho) \\
 &\quad \quad \rho n_1V_1^{-1}V,n_1V_1^{-1} \} V,V_1^{-1}n_1 \\
 &\quad + 2B(1 - B)\rho n_1V_1^{-1}V,n_2V_2^{-1}V,n_1V_1^{-1} \\
 &\quad \quad B^2n_1V_1^{-1}V,n_1V_1^{-1} \\
 &\quad \quad n_1V_1^{-1}\{ 1 - (1 + \rho)(1 - B)^2 \} \\
 &\quad + n_1V_1^{-1}V,n_1V_1^{-1}\{ \rho(1 - B)^2 \\
 &\quad \quad 2B(1 - B)(1 + \rho) - B^2 \} \\
 &\quad + 2\rho B(1 - B)n_1V_1^{-1}V,n_1V_1^{-1}V,n_1V_1^{-1}
 \end{aligned}$$

$$\begin{aligned}
& + 2\rho B(1-B)n_1V_1^{-1}V,n_2V_2^{-1}V,n_1V_1^{-1} \\
& = Bn_1V_1^{-1} + n_1V_1^{-1}V,n_1V_1^{-1}\{\rho/(1+\rho)^2 \\
& - 2\rho/(1+\rho)^2 - \rho^2/(1+\rho)^2\} \\
& = B[n_1V_1^{-1} - n_1V_1^{-1}V,n_1V_1^{-1}]
\end{aligned} \tag{A.7}$$

Analogously, the matrix appearing in $\bar{x}_2^T[\cdot]|\bar{x}_2$ in (4.15) comes out as

$$\begin{aligned}
n_2V_2^{-1} & = \{(1-B)I + Bn_2V_2^{-1}V,\}D_{22}\{(1-B)I + Bn_2V,V_2^{-1}\} \\
& \quad (Bn_2V_2^{-1}V,)D_{11}(Bn_2V,V_2^{-1}) + 2\{(1-B)I + Bn_2V_2^{-1}V,\}D_{12}^T(Bn_2V,V_2^{-1}) \\
& = B[n_2V_2^{-1} - n_2V_2^{-1}V,n_2V_2^{-1}]
\end{aligned} \tag{A.8}$$

We show that the matrix appearing in $\bar{x}_1^T[\cdot]|\bar{x}_2$ in (4.15) is null i.e.,

$$\begin{aligned}
& = \{(1-B)I + Bn_1V_1^{-1}V,\}D_{11}(Bn_2V,V_2^{-1}) \\
& + (Bn_1V_1^{-1}V,)D_{22}\{(1-B)I + BV,n_2V_2^{-1}\} \\
& + 2\{(1-B)I + Bn_1V_1^{-1}V,\}D_{12}\{(1-B)I + Bn_2V,V_2^{-1}\} = 0
\end{aligned} \tag{A.9}$$

Left hand side of (A.9)

$$\begin{aligned}
& = B(1-B)D_{11}V,n_2V_2^{-1} + B(1-B)n_1V_1^{-1}V,D_{22} + 2(1-B)^2D_{12} \\
& + 2B(1-B)(D_{12}V,n_2V_2^{-1} + n_1V_1^{-1}V,D_{12}) \\
& + B^2n_1V_1^{-1}V,(D_{11} + D_{22} - 2D_{12})V,n_2V_2^{-1} \\
& \quad \text{(using (4.8) and (4.16))}
\end{aligned}$$

$$\begin{aligned}
& = B(1-B)\{(1+\rho)n_1V_1^{-1} - \rho n_1V_1^{-1}V,n_1V_1^{-1}\}V,n_2V_2^{-1} \\
& + B(1-B)n_1V_1^{-1}V,\{(1+\rho)n_2V_2^{-1} - \rho n_2V_2^{-1}V,n_2V_2^{-1}\} \\
& + 2(1-B)^2\rho n_1V_1^{-1}V,n_2V_2^{-1} + 2B(1-B)\rho n_1V_1^{-1}V,n_2V_2^{-1}V,n_2V_2^{-1} \\
& + 2B(1-B)\rho n_1V_1^{-1}V,n_1V_1^{-1}V,n_2V_2^{-1} + B^2n_1V_1^{-1}V,n_2V_2^{-1} \\
& \quad \text{(using } V_i^{-1} = n_iV_i^{-1} + n_2V_2^{-1})
\end{aligned}$$

$$= -n_1V_1^{-1}V,n_2V_2^{-1}[-2B(1-B)(1+\rho) + 2\rho(1-B)^2 - B^2 + \rho B(1-B) + 2\rho B(1-B)]$$

0, since $-B = \rho/(1+\rho)$.

Finally, we evaluate the matrix appearing in $\bar{\mathbf{x}}_2^T[\cdot|\bar{\mathbf{x}}_1$ in (4.15), namely

$$\begin{aligned} & -(Bn_2V_2^{-1}V_*)D_{11}\{(1-B)I + BV_*,n_1V_1^{-1}\} - \{(1-B)I \\ & \quad + Bn_2V_2^{-1}V_*\}D_{22}(BV_*,n_1V_1^{-1}) \\ & \quad + 2(Bn_2V_2^{-1}V_*)D_{12}(BV_*,n_1V_1^{-1}) \end{aligned}$$

$$\begin{aligned} & = -B(1-B)n_2V_2^{-1}V_*D_{11} - B(1-B)D_{22}V_*,n_1V_1^{-1} \\ & \quad - B^2n_2V_2^{-1}V_*(D_{11} + D_{22} - 2D_{12})V_*,n_1V_1^{-1} \end{aligned}$$

(using (4.8) and (4.16))

$$\begin{aligned} & = -B(1-B)n_2V_2^{-1}V_*\{(1+\rho)n_1V_1^{-1} - \rho n_1V_1^{-1}V_*,n_1V_1^{-1}\} \\ & \quad - B(1-B)\{(1+\rho)n_2V_2^{-1} - \rho n_2V_2^{-1}V_*,n_2V_2^{-1}\}V_*,n_1V_1^{-1} \\ & \quad - B^2n_2V_2^{-1}V_*,n_1V_1^{-1} \end{aligned}$$

(using $V_*^{-1} = n_1V_1^{-1} + n_2V_2^{-1}$)

$$= -2Bn_2V_2^{-1}V_*,n_1V_1^{-1} \tag{A.10}$$

Combining (A.7)-(A.10), the proof of the lemma is complete.

B. References

- 1 Bancroft, T. A., and Han, C. P. (1981). Inference based on conditionally specified ANOVA models incorporating preliminary testing. *Handbook of Statistics*, VI, Ed. P. R. Krishnaiah. North Holland, New York, 407-441.
- 2 Efron, B., and Morris, C. (1973). Stein's estimation rule and its competitors - an empirical Bayes approach. *J. Amer. Statist. Assoc.*, 68, 117-130.
- 3 Ghosh, M., Saleh, A. K. MD. E., and Sen, P. K. (1987). Empirical Bayes subset estimation in regression models. Tech. Report No. 281. Dept. of Statistics, University of Florida.
- 4 Morris, C. (1983). Parametric empirical Bayes confidence intervals. *Scientific Inference, Data Analysis and Robustness* Ed. G.E.P. Box, T. Leonard and C. F. J. Wu, Academic Press, New York, 25-50.
- 5 Saleh, A. K. Md. E. and Ahmed, S. E. (1987). On preliminary-test and shrinkage estimation of the mean vector of a multivariate normal distribution in a two-sample problem. Private communication.
- 6 Selove, S. L., Morris, C., and Radhakrishnan, R. (1972). Nonoptimality of preliminary test estimators for the multinormal mean. *Ann. Math. Statist.*, 43, 1481-1490.
- 7 Stein, C. (1981). Estimation of the mean of a multivariate normal distribution. *Ann. Statist.*, 9, 1135-1151.
- 8 Strawderman, W. E. (1971). Proper Bayes minimax estimators of the multivariate normal mean. *Ann. Math. Statist.*, 42, 385-388.
- 9 Strawderman, W. E. (1973). Proper Bayes minimax estimators of the multivariate normal mean vector for the case of common unknown variances. *Ann. Statist.*, 1, 1189-1194.

END

DATE

FILMED

5-88

DTIC